

이진 웨이브렛 변환을 이용한 신경회로망의 필기체 문자 인식

이 정 문* · 유 경 산**

A Neural Network Based Handwritten-Character Recognition using Binary Wavelet Transform

Jung-Moon Lee* · Kyoung-San You**

ABSTRACT

In this paper, we propose a new neural pattern recognition from wavelet transform. We first analysis in BFT(Binary Field Transform) in character image. The proposed neural network and wavelet transform is able to improve learning time and scaling. The ability and effectiveness of identifying image using the proposed wavelet transform will be demonstrated by computer simulation.

1. 서 론

컴퓨터의 발달과 더불어 우리는 고도의 정보화 사회로의 진입을 눈앞에 두고 있으며 도서관에서 필요한 책을 찾을 때에도 컴퓨터를 통하여 쉽게 데이터 베이스에서 찾을 수 있다. 근래에 키보드를 대신하여 테블렛이나 스캐너(scanner)를 이용한 문자인식에 관한 연구가 활발하게 진행되고 있다. 숫자인식의 경우에는 필기체로 쓴 우편번호를 사람의 손을 통하지 않고 인식할 수 있다면 많은 시간을 절약할 수 있을 것이다. 문자인식을 통한 응용 분야 중에서 정보의 '자동적인 데이터베이스' 구축을 들 수 있다. 요즈음과 같은 정보화

시대에 매일 수없이 쏟아지는 책을 비롯한 문서들을 문자인식을 통하여 자동적으로 인식시킨다면 사람이 컴퓨터에다 일일이 입력 시키는데 드는 막대한 시간과 인력을 절약할 수 있을 것이다. 문자를 인식하는 방법[4]으로는 원형비교 방법, 통계적 방법, 구조적 방법, 그리고 이외에도 최근 들어 이용되기 시작한 신경회로망을 이용한 방법이 있으며, 본 논문에서 제안하는 기존 신경회로망을 이용한 문자 인식이 아닌 웨이브렛(wavelet)을 이용하여 문자를 분석하여 그 결과 값을 가지고 신경회로망을 이용하여 문자 인식을 하고자 한다.

문자를 인식하거나 도형의 형태를 구분하는 것 등의 시각적인 패턴 인식은 인간에게는 그리 어려운 일이 아니지만 컴퓨터에게는 아직까지도 매우 어려운 일에 속한다. 인간의 두뇌는 100억 개 이상의 매우 많은 뉴런들로 이루어져 있으며, 이들은 서로 밀접한 연관 관계를 가지면서 매우 거대한 네트워크를 형성하기 때문에 매우 복잡한 문제에

* 강원대학교 제어계측공학과 교수

** 강원대학교 제어계측공학과 석사과정

대해서도 뛰어난 해결 능력을 가지고 있다. 비록 인간 두뇌에 있어서 고도의 신경 메카니즘을 규명하기는 매우 어려운 일이지만 두뇌가 시각적인 패턴 인식을 하게 되는 일반적인 아이디어는 신경회로망 응용에 대한 연구를 활성화시키는 동기가 되었다. 시간적, 공간적 데이터에 대한 패턴인식의 신경회로망적인 접근에 있어서의 가장 중요한 이슈는 시각적인 대상을 어떻게 인식할 수 있는냐는 점과 패턴의 시간적인 정보를 어떻게 저장하느냐 하는 점이다. 그리고 또 다른 이슈는 실시간 처리를 위한 데이터의 소규모 작업이다. 클로프(Klopf)와 모건(Morgan)[7]이 지적한대로 “지능에 있어서 실시간 처리는 매우 중요하게 고려되어야 한다”는 점을 간과해서는 안될 것이다. 신경회로망 이론을 통한 시각적 패턴인식[8]중의 하나가 바로 문자 인식이라 할 수 있으며, 문자인식에 이어 영상인식에 대한 관심이 또한 고조되고 있다. 웨이브렛(wavelet)을 이용하여 문자와 영상의 이미지 데이터들을 각각의 데이터의 가장 중요한 특징들만을 선별하여 데이터의 크기를 줄이며 이를 신경회로망의 문자 및 영상 인식에 적용한다. 웨이브렛(wavelet)에서는 문자 데이터들을 2진(binary)으로 만들어서 데이터 처리를 한다. 문자의 가장 기본적인 형태 변형을 압축시켜 본래 문자크기의 4분의 1 수준으로 데이터를 압축시켜 이를 신경회로망의 문자 인식 입력으로 한다. 이것은 불필요한 데이터를 웨이브렛(wavelet)의 전처리 기술을 응용하여 실질적인 문자 본연의 정보만을 추출하여 신경망의 문자 인식 시뮬레이션에 적용하여 신경회로망 연산시 고속 처리가 가능하도록 도와 준다. 본 논문에서는 신경회로망과 웨이브렛(wavelet)의 상관 관계를 응용하여 문자 인식을 한다.

2. Wavelet

본 논문에서 사용하는 Wavelet변환은 2차원 불연속 웨이브렛 변환이며 그림 1 이다.

식 (1) 은 1 차원 신호 $f(x)$ 의 연속 웨이브렛 변환 (Continus Wavelet Transform) [3]이다. a 는 크기, b 는 위치를 나타낸다. 여기서, 식 (2) $\psi_{a,b}$ 는 모(mother) 웨이브렛 ψ 의 전이와 팽창으로 계산되었다. 모(mother) ψ 는 0 의 평균치를 갖는 제약조건이 있다는 가정을 한다.

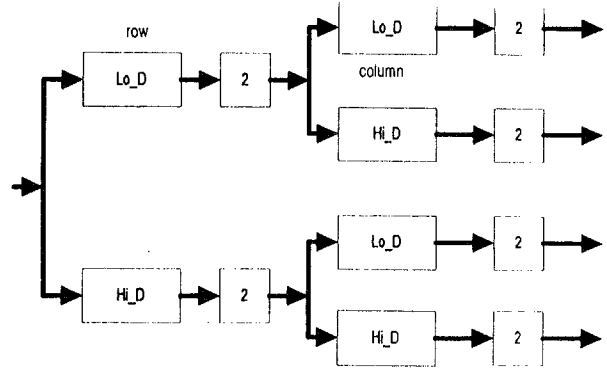


Fig 1. Two-dimensiton DWT

$$(W_a f)(b) = \int f(x) \psi_{a,b}^*(x) dx \quad (1)$$

$$\psi_{a,b}(x) = \frac{1}{\sqrt{a}} \psi\left(\frac{x-b}{a}\right) \quad (2)$$

식 (1) 의 $a=2^n, b \in Z$ 를 이용하여 불연속 웨이브렛 변환을 만들어 낸후 2차원 불연속 웨이브렛 변환 식(3)을 만든다.

$$\begin{aligned} CA_{j+1} &= L_n(b^-) = [H_x * [H_y * L_{n-1}]_{1,2}]_{2,1}(b^-) \\ CD_{j+1}^{(h)} &= D_m(b^-) = [H_x * [G_y * L_{n-1}]_{1,2}]_{2,1}(b^-) \\ CD_{j+1}^{(v)} &= D_{n2}(b^-) = [G_x * [H_y * L_{n-1}]_{1,2}]_{2,1}(b^-) \\ CD_{j+1}^{(d)} &= D_{n3}(b^-) = [G_x * [G_y * L_{n-1}]_{1,2}]_{2,1}(b^-) \end{aligned} \quad (3)$$

$CA_0 = s$ 는 초기치 이다. * 는 컨벌루션 연산자이고, $1 \downarrow 2 (2 \downarrow 1)$ 는 행(렬)의 서브샘플(subsampling)이다. $L_0 = I(x^-)$ 는 초기 원본 이미지(Image) 이다. H 와 G 는 로우패스(low pass) 와 하이 패스(high pass) 의 필터이다. L_n 은 로우패스(low pass)필터로 연산되며 크기 n 의 저 해상도 영상이다. D_{\Rightarrow} 는 특정 방향으로 이루어지는 하이패스(high pass)필터에 의해 연산되고 크기 n 을 가지는 방향에 따른 미세한 정보를 가진다. 초기 원본 이미지 I 는 다양한 크기($\{L_d, D_{\Rightarrow} \mid i=1,2,3, n=1..d\}$)의 집합인 서브이미지(subimage)로 표현된다. a 는 이미지(image) I 의 레벨을 표현하며 본 논문에서는 레벨 1로 연산한다.

3. BFT(Binary Field Transform)

3.1 기본개념

Wavelet base[1][2]를 기반으로 다양한 해상도로 분해하며 선형 시스템형태를 갖는다. binary field 상의 matrix 특성은 real 과 complex field 상의 특징과 다르지만 matrix 연산은 실제로 R and C 의 $GF(2)$ 상에서 수행된다. 유일한 차이점은 모든 연산 2진 연산을 따른다. 견고덕 대문자 A 와 v 는 matrices 와 vectors를 표현한다. A 의 요소 (j, k) 는 $A(j, k)$ 로 표현하고 v 의 요소 k 는 $v(k)$ 로 표현한다. $A(:, j)$ 와 $A(i, :)$ 는 각각 j^{th} 열과 i^{th} 행이다. A^T 는 A 의 transpose 이며, A^{-1} 은 matrix A 의 역행렬이다. 모든 지수는 0부터 시작한다.

$GF(2)$ 의 matrices 와 vectors는 이 논문에서 중요한 특별한 특성을 가진다. 예를 든다면은 vector $v \in R$ or C , $v^T v = 0$ 필요충분조건은 $v = 0$ 이다. $GF(2)$ 의 0이 아닌 짝수 $2n$ 의 숫자는 zero energy를 가진다. (since $0 \equiv 2n \pmod{2}$) $GF(2)$ 의 nonzero vector의 energy는 0이다. $GF(2)$ 의 nonzero element는 1이다. $GF(2)$ 의 $N \times N$ matrix A 는 역으로 $\det(A)=1$ 을 필요충분조건으로 만족한다. 마지막으로 $GF(2)$ 의 단위 matrix U 는 모두 1이 아니다. $M \times N$ matrix U 는 $U^T U = I_N$ 이다. matrix의 각 열은 unit energy를 가지고 있으며, 한쌍의 서로 직교성분을 가지고 있다. 만일 M 이 even 이고, U 의 column이 모두 1이 아니라면 zero energy를 가진다. M 이 odd 이면, U 의 열이 1이 아니라면 다른 모든 열들은 직교조건외 짝수의 1을 가진다. 그러나 이것은 다른 열들은 zero energy를 가진다.

$GF(2)$ 의 4×4 단위 matrix U 의 예가 아래에 있다.

$$U^T U = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix} = I_{4,4} \quad (4)$$

식(4)에서 circulant matrices를 생각 할수 있다.

$N \times N$ matrix $C = [c_{jk}]_{j,k=0,\dots,N-1}$ 은 field F 의 entries를 matrix의 첫 번째 행을 연속적으로 옮긴 것이다. 1만큼 이동하면 1-circulant matrix라고 한다. C 의 one-circulant matrix는 첫 번째 행에 의해 나열된다. 여기는 C 를 $1-circ(c_0, c_1, \dots, c_{N-1})$ 라고 설명한다. 2개의 열을 길이 N 인 한 개의 circular convolution 인 $N \times N$ 의 one-circulant matrix의 하나의 열과 길이 N 인 vector의 다른 열로 표현할수 있다. 인자 m 에 의한 열의 소거는 본래의 열의 모든 m^{th} 개의 샘플에 대한 압축으로 새로운 열을 만들어 낸다. 짝수의 길이 N 인 두 개의 열을 circularly convolve한다면 두 개의 인수에 의한 소거로 convolution 연산이 이루어진다. 그러면 one-circulant matrix와 같이 $\frac{N}{2} \times N$ 인 다음 형식과 같은 two-circulant matrix로 표현되는 소거가 이루어진다.

$$D = \begin{bmatrix} d_0 & d_1 & d_2 & \dots & d_{N-1} \\ d_{N-2} & d_{N-1} & d_0 & \dots & d_{N-3} \\ d_{N-4} & d_{N-3} & d_{N-2} & \dots & d_{N-5} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ d_2 & d_3 & d_4 & \dots & d_1 \end{bmatrix} \quad (5)$$

식(5)는 첫 번째 행에 의해 정해지며 $2-circ(d_0, d_1, \dots, d_{N-1})$ 로 설명되어진다. 다음 절에서 우리가 구상하는 binary wavelet decomposition은 2로 down sampling 되고, wavelet과 scaling function 계수로 이루어진 binary filter의 binary sequences의 circular convolution을 기초로 한다. $GF(2)$ 의 주어진 filtering 연산이 소거를 하고 역을 취했을때의 수열을 결정할수 있는 새로운 이론을 알아 보자. 이 이론을 증명하기 위해서 upper Hessenberg matrix로 transforming circulant matrices를 언급한다. 마치 upper triangular matrix의 계단처럼 보인다. N 이 짝수이고, $N \times N$ upper Hessenberg matrix는 대각선으로 2×2 의 크기의 $\frac{N}{2}$ 의 submatrices를 가진다. 그림 3의 2-band filter bank를 주목하라, 신호는 2개의 필터를 동시에 통과하며 출력은 반으로 감소한다. 필터연산식은 $\frac{N}{2} \times N$ two-circulant matrices 식 (5)의 2개가 조합되어 $N \times N$ matrix를 만들어 낸다.

3.2 Discrete Fourier Transform over GF(2)

비록 DFT(Discrete Fourier Transform)가 real 과 complex 수열에 유용하게 사용된다 하더라도, 이러한 field상태에서는 정확하지는 않다. 좀더 일반적인 정의는 다음과 같다. 필드 F의 임의의 크기 N인 vector v 가 있다. F 상의 v 의 DFT는 다음과 같이 주어진 다른 벡터 \hat{v} 가 있다. 식(6)으로 유도된다.

$$\hat{v}_k = \sum_{i=0}^{N-1} w^{ik} v_i, \quad k=0, \dots, N-1 \quad (6)$$

w는 필드 F의 N차의 원소이다. N차의 $w^N=1$ 과 $w^k \neq 1$ for $k=1, 2, \dots, N-1$ 을 만족한다. DFT 연산은 field F에서 수행되고, 일반적인 모든 특성은 Fourier transform 과 연결된다. 만약 \hat{v} 가 벡터 v의 DFT이면, v는 inverse DFT로 \hat{v} 로부터 복원시킬수 있다. DFT의 유용한 사용으로인해 어떤 fields에서는 N차에서 원소가 없다. 그러므로 이러한 fields에서 N 포인트의 DFT의 값을 구하는 것은 불가능하다. 예를 들어 GF(2)는 $w=1$ 은 0 이 아닌 유일한 원소이다. 즉 원소는 1차 이다. 반면에 GF(2)의 길이 1인 수열의 DFT를 계산할수 있다.

field F에서 N차의 원소가 없을 때, F의 확장 field F^m 에서 크기 N의 DFT를 수행한다. 확장 field는 다항식으로 불리는 구조를 사용해서 만들어진다. m차의 다항식 $p(x)$ 로 이루어지며 F^m 의 모든 원소는 F로부터 얻어지는 계수와 함께 $m-1$ 의 차수를 가지는 다항식이다. 덧셈은 이 새로운 필드에서 다항식의 덧셈이고, 곱셈은 다항식 $p(x)$ 의 다항식의 곱으로 정의되어진다.

이 기술을 응용하여 GF(2)는 m차의 적절한 다항식을 사용하여 GF(2^m)으로 확장되어진다. 그러나, 주어진 N은 크기 N의 DFT 일반 확장 필드는 아니다. 실질적으로 어떤 m에 대해서 짝수의 N의 DFT는 계산할수 없다. 게다가 GF(2^m)에서 원소의 최고차는 2^m-1 이다. 그러므로, 임의의 DFT를 계산하기 위해 임의의 field인 GF(2)를 확장할수 있는 것이다. 예를 들어서, $N=2^m-1$

의 DFT를 계산하기 위해서는 1부터 m까지 데이터 양을 증가시킬 필요가 있고, m차의 다항식을 사용해야만 한다. 실수값을 가지는 신호의 DFT는 N의 독립적인 1부터 2까지의 데이터 차원을 증가시킨다. GF(2)의 고차 다항식의 계산은 transform으로 계산하고 해석하는 것은 매우 어려운 일이다. 소개된 확장 필드의 복잡성을 피하기 위해서 우리는 새로운 binary field transform을 사용할 것이다. GF(2)안의 DFT를 연관시키는 것이 어려운 것은 다른 transform 이 발생하기 때문이다. 예를 들어 GF(2)상의 불연속 Hartley transform 이 길이 N을 가지고 있

으며 DFT로 정의된다고 하자. 그러면 $N=2^m-1$ 과 m으로 분리된다. 실질적으로 N은 짝수가 될 수 없다. 다음의 식(7)은 BFT를 구성하는 수식이며 이는 Figure 1. Walsh waveform을 따른다.

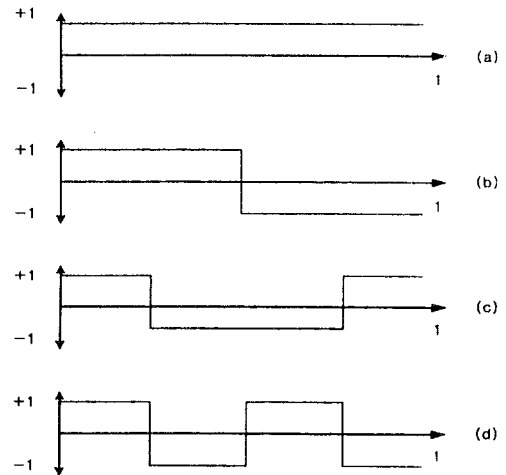


Fig 2 . Walsh transform

$$B_2 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \quad B_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}$$

$$B_N = \begin{bmatrix} B_N^{ul} & B_N^{ur} \\ B_N^{ll} & B_N^{lr} \end{bmatrix}, \quad B_N^{ul} = \begin{bmatrix} 1_{S_{2 \times 2}} & 1_{S_{2 \times (N-4)}} \\ 1_{S_{(N-4) \times 2}} & \bar{B}_{(N-4)} \end{bmatrix}$$

$$B_N^{ur} = \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ 1 & 1 \\ \vdots & \vdots \\ 0 & 0 \end{bmatrix}, B_N^h = B_N^{urT}, B_N^r = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \quad (7)$$

3.3 Filter 구성

BFT 와 GF(2)를 이용하여 필터를 구성한다. 각각의 크기 N인 BFT 행렬식은 대해서 GF(2)^N에서 단위 N인 basis 벡터를 가진다. BFT의 특성상 2진 코드를 사용하기 때문에 연산과정에서 even은 0으로 처리하며 odd는 1로 처리한다. 이때 단위 N의 basis는 이미지의 최소표현이다. BFT의 목적은 이미지를 연산과 표현이 용이한 2진으로 바꾸것이다. 이때 본래 이미지의 정보를 최대한으로 손실시키지 않아야 한다. N=2^K의 스케일링과 BWT(Bianry Wavelet Transform)의 wavelet filter를 설계는 다음과 같다. 먼저 로우패스와 하이패스의 크기 N'을 선택해야 한다. 이는 필터의 범위이다. 이때 이미지 크기를 고려해서 선택해야 한다. 예를 들어 N'=4 이면 이때의 vanishing moments는 2가 된다. 따라서 필터계수는 h₀+h₁=0, h₂+h₃=0이 된다. 필터 h는 2개의 vanishing moments를 가지고 있으므로 h₀+h₂=1, h₁+h₃=1이된다.

h=[h₀ h₀ (1+h₀) (1+h₀)]이 두식의 조합이다. 여기서 나오는 2개의 가능 필터는 [1 1 0 0]과 [0 0 1 1]이 나오게 되면 이것은 h가 N>N' 조건을 만족해야 하며 로우패스 필터에 의해서 결정된다. 로우패스는 BFT의 basis 벡터에 의해서 결정된다. 위의 2.1절의 방법에 적용되는 조건을 따르면은 로우패스는 [1 1 1 0] 과 [1 0 1 1]이 나오게 된다. 이때 두 필터는 N≥4 의 조건을 만족해야 한다.

4. BPN (Back Propagation Neural Network)

4.1 신경회로망

역전파 신경회로망[4]의 학습 알고리즘[5]은 최소자승(least mean square) 알고리즘의 비선형적인

확장이다. 미분의 반복규칙(chain-rule)을 여러번 반복적으로 적용하여 확률 근사치 프레임워크(stochastic - approximation framework)와 관련지음으로써 유도해낼 수 있다.

입력층의 각 유니트에 입력패턴을 주면, 이 신호는 각 유니트에서 변환되어 중간층에 전달되고 최후에 출력층에서 신호를 출력하게 된다. 이 출력값과 기대값을 비교하여 차이를 줄여나가는 방향으로 연결강도를 조절

하고, 상위층에서 역전파하여 하위층에서는 이를 근거로 다시 자기층의 연결강도를 조정해나간다.

학습에서는 입력 및 원하는 출력(목표출력) 패턴(벡터)이 네트워크[6]에 제시된다. 네트워크는 입력층에 주어진 입력패턴이 출력층에 전파되면서 변환 출력패턴을 목표패턴과 비교한다. 네트워크에서 출력된 패턴이 목표패턴과 일치하는 경우에는 학습이 일어나지 않는다. 그렇지 않은 경우는 얻어진 출력패턴과 목표패턴의 차이를 감소시키는 방향으로 네트워크의 연결강도를 조절하여 학습을 한다. 네트워크에 은닉 유니트가 없는 경우는 델타 규칙과 동일하다. 즉 p 번째의 입력/목표출력 패턴이 제시되는 경우에 노드 i에서 노드 j로의 연결강도의 변화는 식 7과 같이 표현된다.

$$\Delta_p w_{ji} = \eta(t_{pj} - o_{pj})i_{i\pi} = \eta \delta_{pj} i_{i\pi} \quad (8)$$

식(8)에서 t_{pj}는 p 번째 목표출력 패턴의 j 성분, o_{pj}는 p 번째 입력패턴으로부터 네트워크가 계산한 출력의 j 성분, i_{iπ}는 p 번째 입력패턴의 i 성분, δ_{pj}=t_{pj}-o_{pj}는 목표출력과 실제출력의 차(오차), Δ_pw_{ji}는 입력층 i 유니트로부터 출력층 j 유니트에의 연결강도의 변화량이다.

4.2 학습신경회로망

본 논문에서 사용될 신경회로망은 그림 3과 같은 신경회로망이다. 입력층으로부터 웨비브렛의 변환을 거쳐 나오는 데이터가 각 특성에 맞게 입력될 수 있도록 4개의 입력 군으로 만들어져 있다. 각각의 LL, LH, HL, LL은 일반적인 신경회로망구조를 가지면 여기서 들어가는 입력 데이터의 특성을 효율적으로 나타내도록 구성되어 있다. 이때 입력 유닛들에 적용되는 Hidden Layer들의 유닛수는 실험적인 결과를 구

할 수밖에 없다. 따라서 입력층의 4개의 군락을 이루는 데이터는 hidden 레이어에서 가중치를 결정하게 되며 가중치들을 모아서 출력층에 전달하게 된다. 일반적인 신경회로망이 모든 유닛에 대해서 가중치를 연결 시키지만 본 논문에서 사용된 신경회로망은 각각의 유닛에 대해서 가중치의 연결을 최소화 하여 계산시간의 단축을 가져오게끔 설계되었다. 기본적인 신경회로망과의 연산 시간과 수정된 신경회로망은 결과에서 나타나듯이 연산시간과 인식능률에 대해서 향상된 결과를 나타낸다.

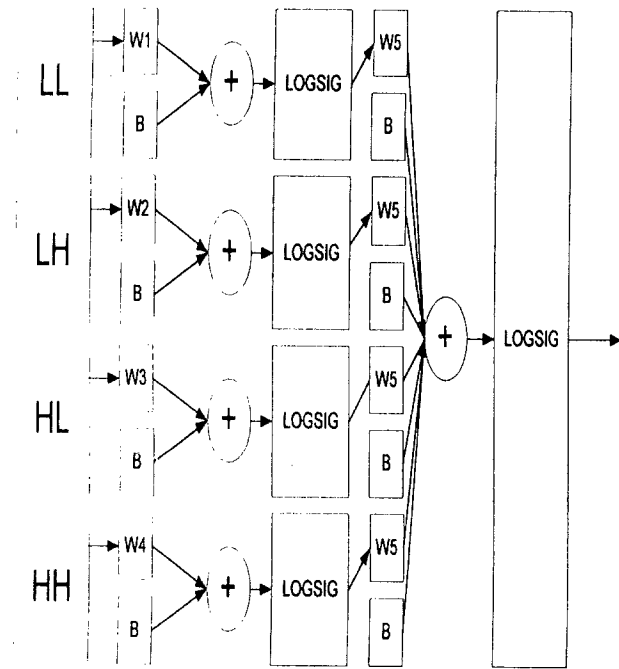


Fig 3 Neural Network

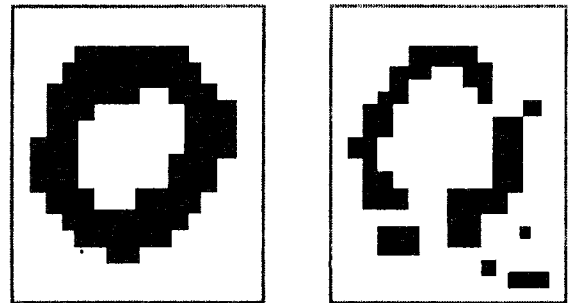
5. 모의 실험

64*4개 유닛의 입력층, 4개의 유닛의 중간층, 10개의 유닛으로 구성되어 있는 출력층으로 네트워크 구조 그림 3 으로 이루어져 있다. 웨이블릿 변환에 있어서 BFT에서 만들어진 l, h가 웨이블릿의 로우 패스와 하이 패스의 필터가 된다. 로우 패스와 하이 패스의 필터는 입력 벡터의 BFT를 이용하여 이미지를 최대한 대로 잘 표현 할수 있는 최적의 필터를 구하는 것이다. 입력된 이미지의 필

0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9

Fig 4. Input Image

터 구성은 Walsh transform으로 구성되어 있다. 2차원 불연속 웨이블릿 변환을 거쳐 나온 웨이블릿 계수가 바로 역전파 신경회로망의 입력값으로 적용된다. 다음 그림 4 는 실제로 역전파 신경회로망에 입력되는 입력값이며, 그림 5는 웨이블릿으로 전처리 하기 위한 이미지 데이터들이다. 그림 5(a)는 그림 4의 이미지들을 바이너리화 하여 새롭게 구성된 바이너리 데이터 들이다. 그리고, 그림 5(b)



(a)

(b)

Fig 5 (a) input data, (b) input with noise

는 이러한 바이너리 이미지들에 대해서 임의의 노이즈를 첨가하여 노이즈를 포함한 데이터에 대한 인식률을 향상하기 위한 데이터이다. 입력에 들어가는 10개의 벡터는 각각에 상응하는 숫자(0..9)를 가지고 있다. 역전파 신경회로망이 정확히 인식하

면 1 이라는 값을 해당 벡터에 출력하게 되며 숫자 인식은 성공적으로 이루어진다. 반면에 해당 벡터를 제외한 나머지 벡터는 모두 0을 출력하여 값이 없게 된다. 역전파 신경회로망은 잡음에 대비하도록 설계 되었으며, 가장 성공적으로 안정하게 인식 하는 범위는 0에서부터 0.2 까지 설계되어 있지만, 같은 네트워크로 웨이브렛의 계수를 사용한 결과 0에서부터 0.5까지의 노이즈에도 인식을 할 수 있다. 네트워크는 2-계층을 가지는 로그 시그모이드 함수를 사용하였으며, 출력의 불린(boolean) 값을 완벽하게 학습하기 위해서 출력 범위를 0에서부터 1까지로 로그 시그모이드 함수에 제한을 두었다. 최초의 학습은 정규화된 데이터로부터 정상적인 학습을 진행하였으며, 학습이 종료된 후 비정규화된 데이터를 적용하여 10번의 반복적인 학습으로 가중치값을 누적 또는 삭감 하였으며, 최종 학습에는 개선된 가중치 값을 이용하여 정규화된 데이터를 학습 시켰다. 입력 데이터는 각각 0-9까지의 샘플데이터 100개를 가지고 있으며 총 1,000개의 데이터를 가진다. 이 중 500개의 데이터로 학습 훈련을 시키고 나머지 500개로 실제 테스트를 한다. 실제로 웨이브렛 변환을 거쳐 나오는 데이터는 8 X 8 이다. 이것은 바로 신경회로망과의 데이터 부족으로 인해 오동작 및 프로그램 수행이 원활하지 않다. 따라서, 새로운 행렬식을 만들어 미완된 부분을 더 보충하는 것이다.

시스템의 성능을 테스트한 결과 노이즈가 섞인 데이터에 대해 일반적인 신경회로망의 인식률과 BFT의 웨이브렛 변환의 전처리 한후의 인식률에 대해 작은 차이를 나타내고 있다. 그림 7은 신경회로망의 인식률을 보여주고 있으며, 다양한 잡음상태에서도 숫자를 인식할수 있음이 확인 되었다. 잡음이 0 인 상태에서부터 0.3 까지 0.05씩 잡음을 증가시켜 비교하여 보았다. 그림 8 은 BFT변환을 한 데이터의 인식률이다.

신경회로망의 2진 이미지의 데이터를 2차 불연속 웨이브렛 변환을 이용하여 각각의 숫자의 특성 계수를 파악하여 역전파 신경회로망의 전처리 기술로 응용한결과 잡음 첨가 숫자 인식에 있어서 0에서부터 0.3 까지의 노이즈에 대해서 인식성능이 안정적인 결과가 나왔다. 반면에 기본적인 역전파 신경회로망으로 학습한 결과는 인식능력에 있어서 웨이브렛으로 변환한 데이터 보다는 인식성능이 떨어진 결과가 나왔다. 신경회로망과 개선된 신경회로망 사이의 수행시간은 입력된 데이터의 노이

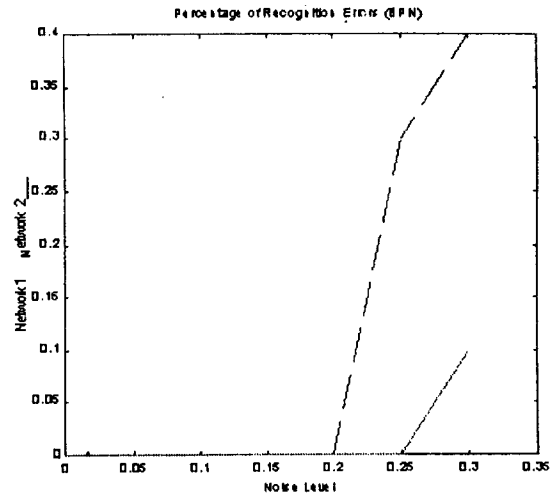


Fig 6 Neural Network Recognition

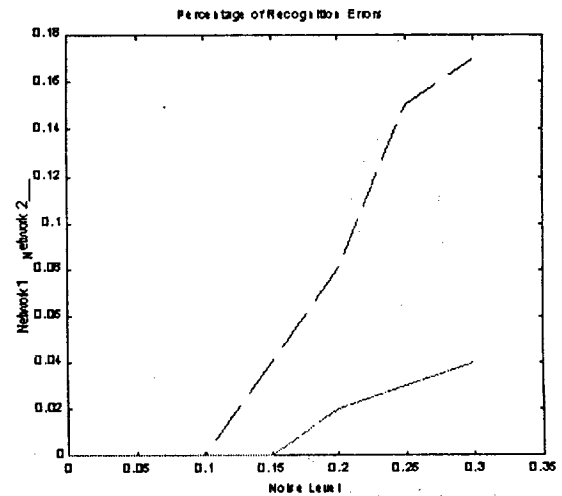


Fig 7 BFT Wavelet Recognition

즈 성분에 따라서 약간의 차이를 보이고 있다. 이는 입력된 노이즈가 랜덤하게 적용되기 때문에 나타나는 현상이지만 평균적인 수치로 확인해 보면 일반 신경회로망은 10-11분의 수행시간이 걸리지만 개선된 신경회로망은 7-8분의 수행시간이 걸린다. 네트워크의 특징상 입력 노드로부터 들어오는 데이터의 크기를 줄인 결과라고 할수 있다. 같은 데이터를 수행하여 시간과 인식률 양쪽으로 성능이 개선됐다.

6. 결 론

같은 역전과 신경회로망을 사용하여 문자 인식을 위한 학습 과정에 있어서 웨이블릿을 이용하여 전처리 한 데이터와 일반적인 데이터 구조를 가진 데이터를 학습한 결과 웨이블릿을 이용하여 학습을 하는 것이 문자 인식에 있어서 인식능력이 우수하다는 결과가 나왔다. 웨이블릿은 각 문자의 고유 특징을 갖고 있으며, 이 특징은 학습 과정에 신경회

로망의 웨이트에 영향을 끼치면서 새로운 데이터 구조를 가지는 것으로 판명 되었다. 이로써 웨이블릿으로 전처리한 데이터로 학습을 하는 것이 신경회로망의 문자인식에 있어 한가지 새로운 방법으로 사용할수 있음을 알 수 있다.

참 고 문 헌

- [1] Zixiang Xiong, "Representation and Coding of Images Using Wavelets", Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Electrical Engineering in the Graduate College of the University of Illinois at Urbana-Champaign, 1996
- [2] P. Scheunders, S. Livens, G. Van de Wouwer, P. vautrot, D.Van Dyck, "Wavelet-based Texture Analysis", *International Journal Computer Science and Information management*, dec. 1997..
- [3] Martin Vetterli, Jelena Kovacevic, "Wavelets and Subband Coding", Prentice hall, Englewood Cliffs, NJ 07632
- [4] Dae Su Kim, "Neural Networks Theory and Applications", Eye Tech Info. 1993
- [5] Stephen T. Welstead, "Neural Network and Fuzzy Logic Applications in C++", John Wiley & Sons, INC.
- [6] James A. Freeman, David M. Skapura, "Neural Networks Algorithms, Applications, and Programming Techniques", Addison-Wesley, 1992
- [7] A. H. Klopff and J. S. Morgan, The role of time in natural intelligence:

Implications for neural network and artificial intelligence research, Proc.

First International Joint Conf. neural Networks, II, pp.97-100, San Diego ,CA, June 18-22, 1989

[8] Kunihiko Fukushima, Neural Network Models for Visual Pattern Recognition, In Parallel Processing in Neural Systems and Computers, pp. 351-356, (Edited by) R. Eckmiller et.al, North-Holland, 1990